PAPER

Loss and Waiting Time Probability Approximation for General Queueing*

Dedicated to Professor N. Suita on his 60th birthday

Kenji NAKAGAWA†, *Member*

SUMMARY Queueing problems are investigated for very wide classes of input traffic and service time models to obtain good loss probability and waiting time probability approximation. The proposed approximation is based on the fundamental recursion formula and the Chernoff bound technique, both of which requires no particular assumption for the stochastic nature of input traffic and service time, such as renewal or markovian properties. The only essential assumption is stationarity. We see that the accuracy of the obtained approximation is confirmed by comparison with computer simulation. There are a number of advantages of the proposed method of approximation when we apply it to network capacity design or path accommodation design problems. First, the proposed method has the advantage of applying to multi-media traffic. In the ATM network, a variety of bursty or non-bursty cell traffic exist and are superposed, so some unified analysis methodology is required without depending each traffic's characteristics. Since our method assumes only the stationarity of input and service process, it is applicable to arbitrary types of cell streams. Further, this approach can be used for the unexpected future traffic models. The second advantage in application is that the proposed probability approximation requires only small amount of computational complexity. Because of the use of the Chernoff bound technique, the convolution of every traffic's probability density fnuction is replaced by the product of probability generating functions. Hence, the proposed method provides a fast algorithm for, say, the call admission control problem. Third, it has the advantage of accuracy. In this paper, we applied the approxmation to the cases of homogeneous CBR traffic, non-homogeneous CBR traffic, $M/D/1$, $AR(1)/D/1$, $M/M/1$ and $D/M/1$. In all cases, the approximating values have enough accuracy for the exact values or computer simulation results from low traffic load to high load. Moreover, in all cases of the numerical comparison, our approximations are upper bounds of the real values. This is very important for the sake of conservative network design.

key words: *loss probability approximation, waiting time probability approximation, general queueing, Chernoff bound*

1. Introduction

Queueing problems are investigated for very wide classes of input traffic and service time models to obtain good loss probability and waiting time probability approximation.

We first consider ATM (Asynchronous Transfer Mode) queueing, i.e., $G/D/1$, because this work was motivated by the analysis of ATM queueing. It is said that in the ATM network, various types of bursty cell streams from many subscribers' terminals are multiplexed and transmitted. Hence, it is not expected that such streams have renewal or markovian properties. So, it is necessary to exploit methods to evaluate network quality without assuming such properties.

We provide cell loss probability approximations in two cases where the stochastic nature of the arriving cell number is given, and next where that of the cell inter-arrival time is given. The goal of our study is to find a good approximation of the tail distribution of the queue length $P[Q>q]$, which can be used as an approximation of the cell loss probability.

Next, we extend our approximating method to general queueing problems. Only the stationarity of customers' arrival process and the service time is assumed. Even the independence of arrival and service process is not required, in general. We provide approximations of $P[Q>q]$ and $P[W>w]$.

The Chernoff bound technique is used to evaluate blocking probabilities in some papers. In [6], the blocking probability at the burst level for the Poisson traffic is approximated by the probability that the amount of input packets exceeds some value. The tail distribution is approximated by the Gaussian distribution, and then the Chernoff bound is applied to obtain an upper bound of the Gaussian tail distribution. In [7], an upper bound of the cell loss probability is calculated for arbitrary traffic that satisfies a given average and peak rate conditions. "The simple burst model" is used as a traffic model and the upper bound of the cell loss probability is shown by the combinatorial expression. Then the upper bound is further approximated by the use of Chernoff bound to reduce the computational complexity.

2. Cell Loss Probability Approximation I of ATM Queueing

We first consider the discrete time ATM queueing system. Every cell which arrives at a single server is

Manuscript received January 20, 1993.
Manuscript revised April 30, 1993.

† The author is with the Faculty of Engineering, Nagaoka University of Technology, Nagaoka-shi, 940-21 Japan.

* A part of this work was done in NTT Transmission Systems Laboratories while the author was a member of the laboratories.

served in a constant service time according to the FIFO (Fist-In-First-Out) discipline. The server is assumed to have a buffer of infinite length. The unit of time is taken to be the service time of the server for one cell. A time interval of the unit length is called a time slot.

Let us denote by a_t the number of cells which arrives at tth time slot and Q_t the queue length at the end of tth time slot. We assume $Q_{t_0}=0$ for some t_0. This assumption is equivalent to the fact that the average number of arriving cells is less than 1. The queue length Q_t satisfies the following well-known recursion formula:

$$Q_t=\max(0, Q_{t-1}-1)+a_t. \tag{1}$$

By solving this recursion, we have the direct expression of Q_t.

Lemma 1: For $t>t_0$, we have

$$Q_t=\max_{0\leq i<t-t_0}\left(\sum_{j=t-i}^{i} a_j-i\right). \tag{2}$$

Proof: Write $Q'_t=Q_t-1$, $a'_t=a_t-1$. Then from (1), we have

$$\begin{aligned}Q'_t&=\max(Q'_{t-1}+a'_t, a'_t)\\&=\max(Q'_{t-2}+a'_{t-1}+a'_t, a'_{t-1}+a'_t, a'_t)\\&=\cdots\\&=\max_{0\leq i<t-t_0}\left(\sum_{j=t-i}^{t} a'_j\right).\end{aligned} \tag{3}$$

We obtain (2) by rewriting (3) with Q_t and a_t.

Lemma 2: If $\{a_t\}$ is a periodic process with period T, then for $t\geqq t_0+T$, we have

$$Q_t=\max_{0\leq i<T}\left(\sum_{j=t-i}^{t} a_j-i\right). \tag{4}$$

Proof: (see [1])

Now, we assume that the arriving cell number $\{a_t\}$ is a stationary process, i.e., the stochastic nature of $\{a_t\}$ is time shift invariant. Let then initial time t_0 tend to $-\infty$, then by (2), we have

$$Q_t=\sup_{i\geqq 0}\left(\sum_{j=t-i}^{t} a_j-i\right). \tag{5}$$

Since the process begins at the infinite past, Q_t and $\sum_{j=t-i}^{t} a_j$ are considered to be in the stationary state. Denote by Q the stationary quene length, and N_i the stationary number of arriving cells in an i-interval. (We call an interval of length i an i-interval.) Thus, by (5), we have

$$Q=\sup_{i\geqq 0}(N_{i+1}-i), \tag{6}$$

and hence

$$P[Q>q]=P[\sup_{i\geqq 1}(N_i-i)\geqq q]. \tag{7}$$

In general, however, it is difficult to calculate the exact value of the right-hand side of (7). In [11], the probabilities of the type as (7) are calculated in the cases where $\{a_t\}$ are i.i.d. or interchangeable random variables.

The goal is to give a good approximation of (7).

First, we provide an upper bound of (7).

Lemma 3: We have

$$P[Q>q]\leqq\sum_{i\geqq 1}P[N_i\geqq i+q] \tag{8}$$

for any $q\geqq 0$.

Proof: Let us define events A and A_i by $A=\{\sup_{i\geqq 1}(N_i-i)\geqq q\}$ and $A_i=\{N_i-i\geqq q\}$, $i\geqq 1$, respectively. Then we have

$$\begin{aligned}P[Q>q]&=P[A]\\&=P\left[\bigcup_{i\geqq 1}A_i\right]\\&\leqq\sum_{i\geqq 1}P[A_i]\\&=\sum_{i\geqq 1}P[N_i\geqq i+q].\end{aligned}$$

In general, let N be a random variable taking on integral values and $P[N=m]$, $m\in\boldsymbol{Z}$, be the probability distribution of N, where $\boldsymbol{Z}$ denotes the set of integers. We define *the generalized probability generating function* $\Psi(z)$ of the random variable N by

$$\Psi(z)\triangleq\sum_{m\in\boldsymbol{Z}}P[N=m]z^m, \tag{9}$$

if the infinite series of the right-hand side of (9) converges. If N takes only on non-negative integral values, the generalized probability generating function is the usual probability generating function.

In order to estimate each term of the right-hand side of (8), we use the Chernoff bound technique.

Lemma 4: (Chernoff bound) Let N be a random variable taking on integral values and $\Psi(z)$ the generalized probability generating function of N. Then

$$P[N\geqq r]\leqq\alpha^{-r}\Psi(\alpha) \tag{10}$$

holds for any real number r and α, $\alpha\geqq 1$.

Proof: By Markov's inequality, for $\alpha>1$ we have

$$\begin{aligned}P[N\geqq r]&=P[\alpha^N\geqq\alpha^r]\\&\leqq\alpha^{-r}E[\alpha^N]\\&=\alpha^{-r}\Psi(\alpha).\end{aligned}$$

The upper bound of (10) is the tightest if α is the value attaining the minimum of $\alpha^{-r}\Psi(\alpha)$, $\alpha\geqq 1$.

By summarizing above, we have

Theorem 1: Let $\Psi_i(z)$ be the probability generating function of the random variable N_i, $i=1, 2, \cdots$. Then, we have

$$P[Q>q]\leqq u_1(q), \tag{11}$$

where

$$u_1(q)=\sum_{i\geq 1} \alpha_i^{-(i+q)} \Psi_i(\alpha_i), \tag{12}$$

and α_i is the number that minimizes $\alpha^{-(i+q)}\Psi_i(\alpha)$, $\alpha\geqq 1$, $i=1, 2, \cdots$.

2.1 Application

2.1.1 CBR Traffic with Different Periods

We consider the cell loss probability of superposed CBR (Continuous Bit-Rate) traffic. The cell inter-arrival time distribution of a CBR trunk is the deterministic distribution. The length between two adjacent cells is called the period of the CBR. The queueing system of superposed CBR traffic with different periods is represented by $\sum D_i/D/1$, where D_i denotes the deterministic distribution of arbitrary cell inter-arrival time. We consider the superposition of CBR with n different periods $M_j, j=1, \cdots, n$. The number of input trunks with period M_j is denoted by $K_j, j=1, \cdots, n$, and the total number of trunks $K=\sum_{j=1}^{n} K_j$. Then the link utilization ρ is given by $\rho=\sum_{j=1}^{n} K_j/M_j$.

Many authors have investigated the calculation of the cell loss probability for CBR traffic [1]-[4], [8], [9], [12]. In [1]-[4] and [12], the authors considered the superposition of CBR cell streams with a single cell period, which can be represented by $KD/D/1$. They obtained exact formulas of $P[Q>q]$, some of which are direct formulas [1], [2], and the others are recursions [3], [4], [12].

In [2], an asymptotic method is used to approximate $P[Q>q]$. The obtained formula is a very simple function of q, and therefore the computational time complexity is practically on the order of constant. That asymptotic method gives a good approximation if the link utilization is high, but otherwise it does not. Unfortunately, the formula in [2] gives a lower estimate for $P[Q>q]$ in many cases, hence it is not appropriate to use it for designing network system parameters.

It is not possible to apply the above-mentioned formulas ([1]-[4], [12]) to the case of superposing CBR traffic with different periods.

In [9], on the other hand, the cell loss probabilities for CBR with different periods are studied and upper and lower bounds of $P[Q>q]$ are obtained. From the viewpoint of computational complexity; however, it is difficult to apply this result when the number of different periods is large. Because in order to obtain the upper bound by the method of [9], the probability density functions of arriving cell numbers on all the CBR trunks should be convoluted. This requires considerable computational complexity.

In the case of CBR with different periods, exact calculation formulas of $P[Q>q]$ are not known except for very limited cases [8].

Now, in order to apply our Theorem 1 to the calculation of the cell loss probability approximation of $\sum D_i/D/1$ queueing, the probability generating functions of cell arriving numbers should be calculated.

Let $N_{ij}^{(1)}$ denote the number of arriving cells from a single CBR trunk with period M_j during an i-interval, and $\Psi_{ij}^{(1)}(z)$ the probability generating function of $N_{ij}^{(1)}$, $i=1, 2, \cdots, j=1, \cdots, n$, i.e.,

$$\Psi_{ij}^{(1)}(z)=\sum_{m=0}^{\infty} P[N_{ij}^{(1)}=m]z^m. \tag{13}$$

In this case, $\Psi_{ij}^{(1)}(z)$ is easily calculated:

$$\Psi_{ij}^{(1)}(z)=\left(\frac{i}{M_j}-\left\lfloor\frac{i}{M_j}\right\rfloor\right)z^{\lfloor\frac{i}{M_j}\rfloor+1} +\left(1-\frac{i}{M_j}+\left\lfloor\frac{i}{M_j}\right\rfloor\right)z^{\lfloor\frac{i}{M_j}\rfloor+1}, \tag{14}$$

where $\lfloor x \rfloor$ denotes the maximum integer not greater than x. The probability generating function $\Psi_i(z)$ of N_i is given by the product of all $\Psi_{ij}^{(1)}(z)$:

$$\Psi_i(z)=\prod_{j=1}^{n} \Psi_{ij}^{(1)}(z)^{K_j}. \tag{15}$$

The upper bound $u_1(q)$ is obtained by substituting (15) into (12).

Particularly, let us consider the homogeneous CBR traffic with a single period M and the number of trunks K. The condition $K<M$ is assumed for system's stability. By Lemma 2, we have

$$Q=\max_{0\leq i<M} (N_{i+1}-i), \tag{16}$$

and then

$$P[Q>q]=P\left[\max_{1\leq i\leq M} (N_i-i)\geqq q\right]. \tag{17}$$

Since $N_i\leqq K$, $i=1, \cdots, M$, if $i>K-q$ we have $N_i-i<q$. Thus the maximum in (17) is not attained by $i>K-q$. So, in the homogeneous CBR case, we have

$$P[Q>q]=P\left[\max_{1\leq i\leq K-q} (N_i-i)\geqq q\right] \tag{18}$$

$$\leqq \sum_{i=1}^{K-q} P[N_i\geqq i+q]. \tag{19}$$

The probability generating function $\Psi_i(z)$ of N_i is

$$\Psi_i(z)=\left(\frac{i}{M}z+1-\frac{i}{M}\right)^K, \quad i=1, \cdots, K-q. \tag{20}$$

Hence by Theorem 1, the upper bound $u_1(q)$ in the homogeneous case is explicitly written as

$$u_i(q)=\left(\frac{K}{M}\right)^K \sum_{i=1}^{K-q}\left(\frac{M-i}{K-q-i}\right)^{K-q-i}\left(\frac{i}{q+i}\right)^{q+i}, \tag{21}$$

where $(M-K+q/0)^0$ is assumed to imply 1.

2.1.2 $M/D/1$

Let us consider $M/D/1$ queueing system. The probability generating function $\Psi_i(z)$ of the arriving cell number N_i during an i-interval of the Poisson traffic at the mean arrival rate ρ is given by

$$\Psi_i(z)=e^{\rho i(z-1)}, \quad i=1,2,\cdots. \tag{22}$$

Hence, we have the upper bound $u_1(q)$ as

$$u_1(q)=\sum_{i\geq 1}\left(\frac{\rho i}{i+q}\right)^{i+q}e^{i+q-\rho i} \tag{23}$$

for any integer $q>0$.

The exact survivor function $P[Q>q]$ of $M/D/1$ queueing is known [9] as

$$P[Q>q]=1-(1-\rho)\sum_{j=0}^{q}\frac{(-\rho j)^{q-j}}{(q-j)!}\exp(\rho j). \tag{24}$$

We will later give numerical comparison of our approximation and the formula (24).

Remark: The formula (24) is not suitable for stable calculation because positive and negative numbers of large absolute values are added alternatively. In the real computation, the following stably computable formula (see [5]) is used;

$$P[Q>q]=1-\sum_{j=1}^{q}p_j, \tag{25}$$

where $p_j=P[Q=j]$, which is calculated recursively by

$$p_0=1-\rho, \quad p_1=(1-\rho)(-1+e^{\rho}),$$

$$p_j=\sum_{k=1}^{j}e^{\rho}\frac{(-\rho)^{k-1}}{(k-1)!}p_{j-k}, \quad j\geq 2. \tag{26}$$

2.1.3 $AR(1)/D/1$

Next, we consider the case where the input process is represented by a time series analytic model. Suppose the input process $\{a_t\}$ is represented as

$$a_t=\rho+\xi_t,$$

$$\xi_t=b\xi_{t-1}+\varepsilon_t, \tag{27}$$

where ρ is the mean arrival rate of $\{a_t\}$ and b is a constant, $|b|<1$, and $\{\varepsilon_t\}$ are i.i.d. Gaussian random variables whose means are 0 and the variances are σ^2 i.e., $\varepsilon_t\cdots\mathcal{N}(0,\sigma^2)$. The process $\{\xi_t\}$ defined by (27) is called an $AR(1)$ process, or autoregressive process of order 1. It is known that $\{\xi_t\}$ has the stationary representation

$$\xi_t=\sum_{j=1}^{\infty}b^j\varepsilon_{t-j}, \tag{28}$$

(see, for example, [10], pp. 391).

Hence, by calculation, we see that the mean of N_i is ρi, and the variance σ_i^2 of N_i is

$$\sigma_i^2=\mathrm{Var}(N_i)$$

$$=E\left[\left(\sum_{t=1}^{i}\sum_{j=0}^{\infty}b^j\varepsilon_{t-j}\right)^2\right]$$

$$=\frac{\sigma^2}{1-b^2}\frac{1}{(1-b)^2}((1-b^2)i-2b(1-b^i)).$$

Hence, we see that N_i satisfies $N_i\cdots\mathcal{N}(\rho i,\sigma_i^2)$.

In general, for a continuous random variable X taking on real values with probability density function $p(x)$, *the generalized probability generating function* $\Psi_X(z)$ of X is defined by

$$\Psi_X(z)=E[z^X]=\int_{-\infty}^{\infty}z^x p(x)\,dx. \tag{29}$$

By slightly modifying the proof, we see that the Theorem 1 holds for continuous random variables. Therefore, we have the upper bound $u_1(q)$ of $P[Q>q]$ of $AR(1)/D/1$ as follows:

$$P[Q>q]\leq u_1(q)=\sum_{i\geq 1}e^{-(q+i-\rho i)^2/2\sigma_i^2}. \tag{30}$$

3. Cell Loss Probability Approximation II of ATM Queueing

In this section, we provide an upper bound of $P[Q>q]$ in the ATM queueing system when the cell inter-arrival time distribution is given.

Let Q_n denote the number of cells that are already in the buffer when the nth cell arrived, and t_n the arrival time of the nth cell. Denote by $\tau_n=t_n-t_{n-1}$ the inter-arrival time between nth and $n-1$th cells. We have the recursion formula of $\{Q_n\}$;

$$Q_n=\max(0, Q_{n-1}+1-\tau_n). \tag{31}$$

Let us denote by Q the limit of Q_n as n tends to infinity, and $U_n=\sum_{i=0}^{n-1}(1-\tau_{i+1})=n-\sum_{i=1}^{n}\tau_i$, $n>0$, $U_0=0$. By solving (31), we have the direct expression of Q (see [5]);

$$Q=\sup_{n\geq 0}U_n. \tag{32}$$

Then, we have an upper bound of $P[Q>q]$.

Theorem 2: Write $\Psi_n(z)$ the probability generating function of the random variable $\sum_{i=1}^{n}\tau_i$, then we have

$$P[Q>q]\leq u_2(q), \tag{33}$$

where

$$u_2(q)=\sum_{n>q}\alpha_n^{n-q}\Psi_n(\alpha_n^{-1}), \tag{34}$$

and α_n is the value of α that minimizes $\alpha^{n-q}\Psi_n(\alpha^{-1})$, $n>q$.

Proof: From (32), we have

$$P[Q>q]=P\left[\sup_{n\geq 0}\left(n-\sum_{i=1}^{n}\tau_i\right)>q\right]$$

$$=P\left[\sup_{n>q}\left(n-\sum_{i=1}^{n}\tau_i\right)>q\right]$$
$$\leqq \sum_{n>q} P\left[n-\sum_{i=1}^{n}\tau_i>q\right]$$
$$\leqq \sum_{n>q} \alpha_n^{n-q}\Psi_n(\alpha_n^{-1}), \qquad (35)$$

Notice that the generalized probability generating function of $-\sum\tau_i$ is $\Psi_n(z^{-1})$.

3.1 Application to $M/D/1$

We apply Theorem 2 to $M/D/1$ queueing where the cell inter-arrival time is described by the exponential distribution. The probability generating function $\Psi(z)$ of the exponential distribution at mean $1/\rho$ is $\Psi(z)=\rho/(\rho-\log z)$ and that of $\sum_{i=1}^{n}\tau_i$ is $\Psi_n(z)=\Psi^n(z)=(\rho/(\rho-\log z))^n$. Hence, we have the upper bound $u_2(q)$ of $M/D/1$ cell loss probability as follows:

$$u_2(q)=\sum_{n>q} e^{n-\rho(n-q)}\left(\frac{\rho(n-q)}{n}\right)^n. \qquad (36)$$

We find that the upper bound (36) is the same as (23).

4. Loss Probability Approximation of $G/G/1$

We extend the above approximation methods to general queueing problems, i.e., $G/G/1$.

We first study the case where the input traffic and the service time are both stationary and mutually independent. The independence assumption is not essential, however, it is assumed in this section, because the independent case gives us simple calculation results. We will later discuss the general case (non-independent case).

Let us denote by t the discrete time. Customers are served in the FIFO discipline by a single server with a infinite length waiting room. Denote by a_t, b_t, and Q_t the number of arriving customers at tth time slot, the number of customers served at tth time slot, and the queue length at tth time slot, respectively. (First, customers arrive, second, queue length is counted, and then customers are served.) We assume $Q_{t_0}=0$ for some t_0. Then we have the fundamental recursion

$$Q_t=\max(0, Q_{t-1}-b_{t-1})+a_t. \qquad (37)$$

The recursion (37) leads to the direct expression;

$$Q_t=\max_{0\leqq i<t-t_0}\left(\sum_{j=t-i}^{t} a_j-\sum_{j=t-i}^{t-1} b_j\right). \qquad (38)$$

The empty sum in (38) is assumed to imply 0. Let us denote by Q the stationary number of customers in the queue, N_i the stationary number of arriving customers during an i-interval and L_i the stationary number of customers served during an i-interval. Further, denote by $\Psi_{N_i}(z)$ and $\Psi_{L_i}(z)$ the probability generating function of N_i and L_i, respectively. Then we have the following equation in the same way as (7).

$$Q=\sup_{i\geqq 1}(N_i-L_{i-1}), \quad L_0=0 \qquad (39)$$

Hence, we have

Theorem 3: For stationary input traffic and service time which are mutually independent, the loss probability is bounded by

$$P[Q>q]\leqq u_3(q). \qquad (40)$$

where

$$u_3(q)=\sum_{i\geqq 1}\alpha_i^{-(q+1)}\Psi_{N_i}(\alpha_i)\,\Psi_{L_{i-1}}(\alpha_i^{-1}), \qquad (41)$$

and $\Psi_{N_i}(z)$ and $\Psi_{L_i}(z)$ are the probability generating functions of the stationary number N_i of arriving customers and the stationary number L_i of served customers during i-interval, respectively. Further, α_i is the number that minimizes $\alpha^{-(q+1)}\Psi_{N_i}(\alpha)\,\Psi_{L_{i-1}}(\alpha^{-1})$, $\alpha\geqq 1$, $i=1, 2, \cdots$.

Moreover, we can eliminate the independence assumption of the input traffic and the service time. If the joint distribution of the input and the service time is given, an upper bound can be obtained in the same way as above with Chernoff bound technique.

4.1 Application to $M/M/1$

Let us apply Theorem 3 to $M/M/1$ queueing. The numbers of both arriving and served customers are Poisson, thus we have $\Psi_{N_i}(z)=e^{\rho i(z-1)}$ and $\Psi_{L_{i-1}}(z^{-1})=e^{\mu(i-1)(z^{-1}-1)}$, where ρ and μ are the mean values of the two Poisson distributions, respectively.

Then the loss probability of $M/M/1$ is bounded by

$$P[Q>q]\leqq u_3(q)=\sum_{i\geqq 1}\alpha_i^{-(q+1)}e^{\rho i(\alpha_i-1)+\mu(i-1)(\alpha_i^{-1}-1)}, \qquad (42)$$

where α_i is the positive root of the equation $\rho i\alpha^2-(q+1)\alpha-\mu(i-1)=0$.

We will compare this bound with the exact value $P[Q>q]$ of $M/M/1$, that is,

$$P[Q>q]=\left(\frac{\rho}{\mu}\right)^{q+1}. \qquad (43)$$

5. Waiting Time Probability Approximation of $G/G/1$

We consider here the waiting time of a customer of $G/G/1$ queueing system.

Let us denote by σ_n the service time of the nth customer, τ_n the inter-arrival time between nth and

$n-1$th customers, and W_n the waiting time of the nth customer. Then we have the fundamental recursion formula:

$$W_n=\max(0,\ W_{n-1}+\sigma_{n-1}-\tau_n). \quad (44)$$

We assume that the service time and the inter-arrival time are stationary and mutually independent. Write $V_n=\sum_{i=1}^{n}(\sigma_{i-1}-\tau_i)$, $n>0$, $V_0=0$, then we have the direct expression of the stationary waiting time W [5];

$$W=\sup_{n\geqq 0} V_n. \quad (45)$$

Denote by S_n the stationary service time for consecutive n customers, T_n the stationary inter-arrival time between kth and $k+n$th customers. Then, in the same way as the previous discussion, we have

$$P[W>w]=P[\sup_{n\geqq 1} V_n>w]$$
$$\leqq \sum_{n\geqq 1} P[S_n-T_n>w].$$

We write $\Psi_n^{\sigma}(z)$, $\Psi_n^{\tau}(z)$ the probability generating functions of S_n, T_n, respectively. Then we have

Theorem 4: For stationary service time and inter-arrival time which are mutually independent, the probability that the waiting time is more than w is bounded as follows.

$$P[W>w]\leqq u_4(w), \quad (46)$$

where

$$u_4(w)=\sum_{n\geqq 1} a_n^{-w}\Psi_n^{\sigma}(a_n)\Psi_n^{\tau}(a_n^{-1}), \quad (47)$$

and $\Psi_n^{\sigma}(z)$ is the probability generating function of the stationary service time S_n for consecutive n customers, $\Psi_n^{\tau}(z)$ is that of the stationary inter-arrival time T_n between kth and $k+n$th customers. Further, a_n is the value of a that minimizes $a^{-w}\Psi_n^{\sigma}(a)\Psi_n^{\tau}(a^{-1})$, $a\geqq 1$, $n=0, 1, \cdots$.

5.1 Application to $D/M/1$

Let us consider the $D/M/1$ queueing. Denote by ρ the mean arrival rate and μ the mean service rate. In this case, we have $\Psi_n^{\tau}(z)=z^{n/\rho}$ and $\Psi_n^{\sigma}(z)=(\mu/(\mu-\log z))^n$. Hence the upper bound $u_4(w)$ is

$$u_4(w)=\sum_{n\geqq 1}\left(\frac{\mu}{\rho}+\frac{\mu}{n}w\right)^n e^{-\mu w+(1-\mu/\rho)n}. \quad (48)$$

The exact formula of $P[W>w]$ of $D/M/1$ queueing is given [5] by

$$P[W>w]=\eta^{1+\rho w}, \quad (49)$$

where η is the unique solution of the equation

$$\eta=e^{-\mu(1-\eta)/\rho},\quad 0<\eta<1. \quad (50)$$

6. Heuristic Modification

From detailed and extensive numerical comparison between obtained upper bounds and the exact formulas or simulation, it seems that $\log P[Q>q]$ is approximated well by $\log u_k(q)+$constant, $k=1, 2, 3$. Similarly, $\log P[W>w]$ is approximated well by $\log u_4(w)+$constant. Thus we here make the following modification:

$$\tilde{u}_k(q)=C_0 u_k(q),\quad k=1, 2, 3,$$
$$\tilde{u}_4(w)=C_0 u_4(w). \quad (51)$$

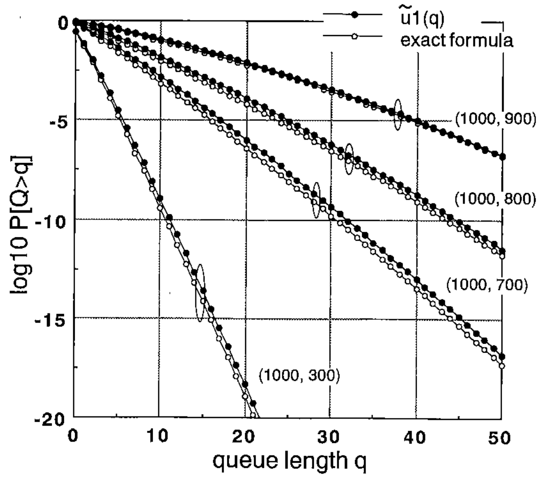


Fig. 1 Homogeneous CBR traffic; Comparison between $\tilde{u}_1(q)=C_0\times(21)$ and Bhargava's exact formula (52) with parameters (M, K), where M is the period of a CBR trunk and K is the number of trunks.

where $C_0=P[Q>0]/u_k(0)$, $k=1, 2, 3$ or $C_0=P[W>0]/u_4(0)$. We show the numerical calculation results of $\tilde{u}_k(q)$ and $\tilde{u}_4(w)$ in Figs. 1-7 to compare with the exact formulas or computer simulation.

7. Numerical Results

In this section, the obtained approximation $\tilde{u}_i(q)$, $i=1, \cdots, 3$ and $\tilde{u}_4(w)$ are compared numerically with $P[Q>q]$ or $P[W>w]$ for the queueing systems treated in the above sections.

In Fig. 1, we show the comparison of the approximation $\tilde{u}_1(q)=C_0\times(21)$ and the exact formula of the cell loss probability of the superposed homogeneous CBR traffic with period $M=1000$ and the number of input trunks $K=300$, 700, 800 and 900. The exact formula of the homogeneous CBR given by Bhargava [1] is

$$P[Q>q]=\sum_{j=1}^{K-q}\frac{M-K+q}{M-j}\binom{K}{j+q}$$

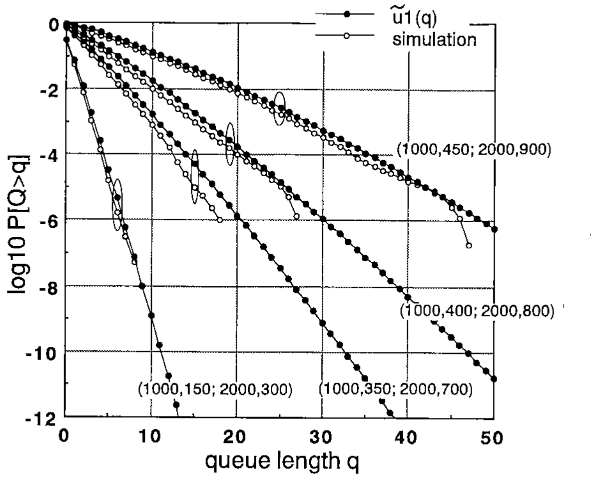


Fig. 2 Non-homogeneous CBR traffic; Comparison between $\tilde{u}_1(q)$ and the computer simulation with parameters $(M_1, K_1; M_2, K_2)$, where M_i is the period and K_i is the number of trunks with period M_i, $i=1, 2$.

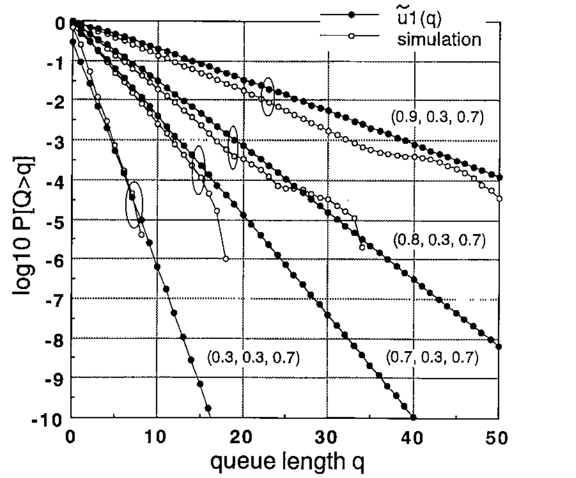


Fig. 4 $AB(1)/D/1$; Comparison between $\tilde{u}_1(q)=C_0\times(30)$ and the computer simulation with parameters (ρ, b, σ), where ρ is the mean arrival rate, b the AR coefficient, and σ the standard deviation of the innovation process. $(b=0.3, \sigma=0.7)$.

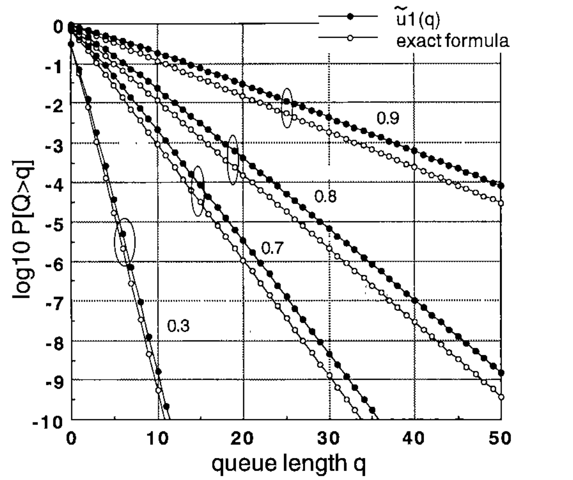


Fig. 3 $M/D/1$; Comparison between $\tilde{u}_1(q)=C_0\times(23)$ and the exact formula (24) with parameter ρ, where ρ is the mean arrival rate of Poisson.

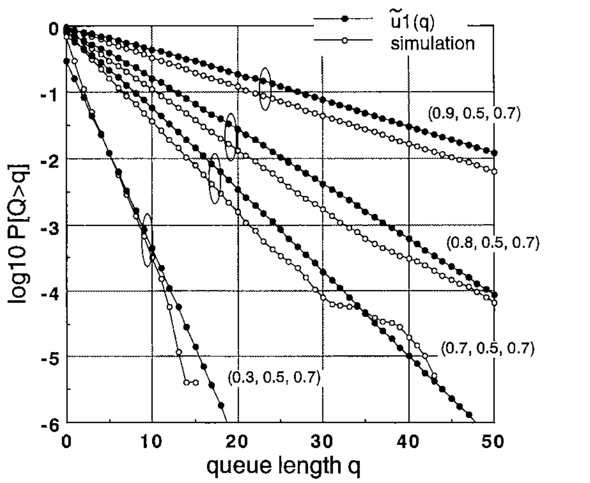


Fig. 5 $AR(1)/D/1$; Comparison between $\tilde{u}_1(q)=C_0\times(30)$ and the computer simulation with parameters (ρ, b, σ), where ρ is the mean arrival rate, b the AR coefficient, and σ the standard deviation of the innovation process. $(b=0.5, \sigma=0.7)$.

$$\cdot\left(\frac{j}{M}\right)^{j+q}\left(1-\frac{j}{M}\right)^{K-j-q} \qquad (52)$$

In Fig. 2, we show the comparison of the approximation $\tilde{u}_1(q)$ and the simulation result of superposed non-homogeneous CBR traffic with two different periods $M_1=1000$, $M_2=2000$ and the number of trunks $K_1=150$, 350, 400 and 450, $K_2=300$, 700, 800 and 900.

In Fig. 3, we show $\tilde{u}_1(q)=C_0\times(23)$ and the exact formula (24) of $M/D/1$ with the mean arrival rates $\rho=0.3$, 0.7, 0.8 and 0.9.

In Figs. 4 and 5, we show the results of the $AR(1)/D/1$ queueing. The approximation $\tilde{u}_1(q)=C_0\times(30)$ and the simulation results are shown in the cases where the mean arrival rates $\rho=0.3$, 0.7, 0.8 and 0.9, and the AR coefficients $b=0.3$ and 0.7. The standard deviation σ of the innovation process is $\sigma=0.7$.

In Fig. 6, we show the approximation $\tilde{u}_3(q)=C_0\times(42)$ and the $M/M/1$ exact formula (43) with the mean service rate $\mu=1$ and the mean arrival rates $\rho=0.3$, 0.7, 0.8 and 0.9.

In Fig. 7, we show the comparison of $\tilde{u}_4(w)=C_0\times(48)$ and the $D/M/1$ exact formula (49) in the cases $\mu=1$ and $\rho=0.3$, 0.7, 0.8 and 0.9.

In all cases, we see that $\tilde{u}_k(q)$ $\tilde{u}_4(w)$ approximate well the exact value or the simulation results of $P[Q>q]$ or $P[W>w]$.

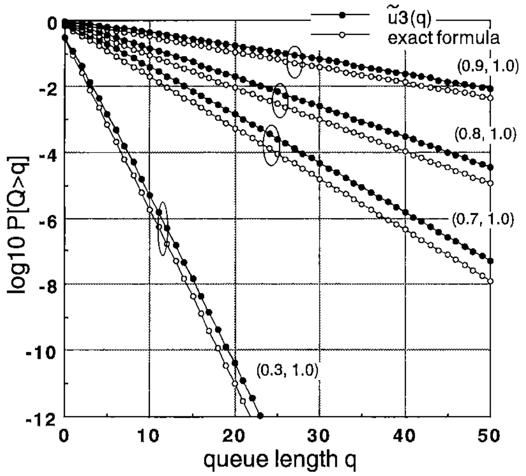


Fig. 6 $M/M/1$; Comparison between $\tilde{u}_3(q)=C_0\times(42)$ and the exact formula (43) with parameters (ρ, μ), where ρ is the mean arrival rate and μ is the mean service rate.

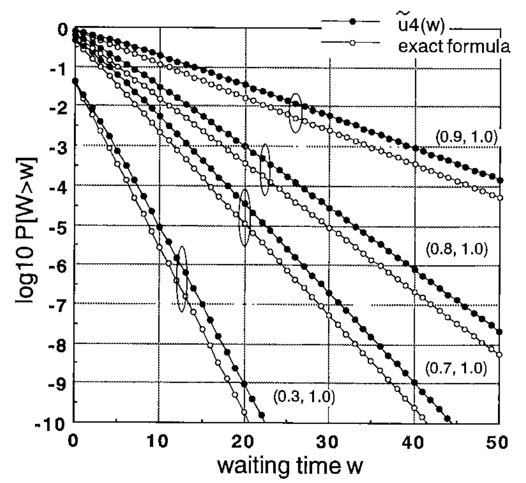


Fig. 7 $D/M/1$; Comparison between $\tilde{u}_4(w)=C_0\times(48)$ and the exact formula (49) with parameters (ρ, μ), where ρ is the mean arrival rate and μ is the mean service rate.

8. Conclusion

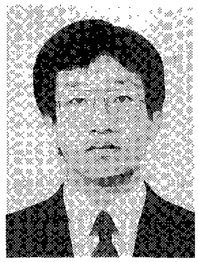
Four kinds of approximation are presented for the loss survivor function $P[Q>q]$ or the waiting time survivor function $P[W>w]$ of general queueing problems. The assumption on the input traffic and the service time process is only the stationarity, even the independence of the input and the service is not assumed.

We showed the comparison between our approximation and the exact formula or the simulation results of $P[Q>q]$ or $P[W>w]$ in the cases of homogeneous CBR traffic, non-homogeneous CBR traffic, $M/D/1$, $AR(1)/D/1$, $M/M/1$ and $D/M/1$. We confirmed that, in all cases, $\tilde{u}_k(q)$ and $\tilde{u}_4(w)$ approximate the real values well.

One of the other advantages of our approximating method is that it is applicable to arbitrary type of independent superposition of input traffic, such as $D+M+AR(1)/G/1$, and so forth. For, the probability generating function in the expression of $\tilde{u}_k(q)$ is obtained by the product of all input traffic's probability generating functions. Hence, it is widely applicable to the analysis of so called multi media traffic. In a further study, we will apply $\tilde{u}_k(q)$ and $\tilde{u}_4(w)$ to wider classes of input and service time processes.

References

[1] Bhargava, A., Humblet, P. and Hluchyj, M. G., "Queueing Analysis of Continuous Bit-stream Transport in Packet Networks," in *Proc. IEEE GLOBE-COM*, pp. 930-907, 1989.
[2] Dron, L. G., Ramamurthy, G. and Sengupta, B., "Delay Analysis of Continuous Bit Rate Traffic Over an ATM Network," *IEEE J. Sel. Areas Commun.*, vol. 9, no. 3, pp. 402-407, Apr. 1991.
[3] Eckberg, A. E., Jr., "The Single Server Queue with Periodic Arrival Process and Deterministic Service Times," *IEEE Trans. Commun.* vol. 27, no. 3, pp. 556-562, Mar. 1979.
[4] Karol, M. J. and Hluchyj, M. G., "Using a Packet Switch for Circuit-Switched Traffic: A Queueing System with Periodic Input Traffic," *ICC'87*, Seattle, pp. 1677-1682.
[5] Kleinrock, L., *Queueing Systems, I, II*, Wiley, New York, 1975.
[6] Hui, J. Y., "Resource Allocation for Broadband Networks," *IEEE J.* Sel. Areas Commun., vol. 6, no. 9, pp. 1598-1608, Dec. 1988.
[7] Miyao, Y., "A Call Admission Control Scheme in ATM Networks," *ICC'91*, pp. 391-396.
[8] Nakagwa, K., unpublished work
[9] Roberts, J. W. and Virtamo, J. T., "The Superposition of Periodic Cell Arrival Streams in an ATM Multiplexer," *IEEE Trans. Commun.*, vol. 39, no. 2, pp. 298-303, Feb. 1991.
[10] Shiryayev, A. N., *Probability*, Springer-Verlag, 1984.
[11] Takács, L., *Combinatorial Methods in the Theory of Stochastic Processes*, Wiley, New York, 1967.
[12] Wong, A. K., "Queueing Analysis for ATM Switching of Continuous-Bit-Rate Traffic — A Recursion Computation Method," in *Proc. IEEE GLOBECOM*, pp. 1438-1444, 1990.



Kenji Nakagawa was born in Yamanashi, Japan, on 28 January, 1957. He received the B.S., M.S. and D.S. degrees in mathematics from Tokyo Institute of Technology in 1980, 1982, and 1986, respectively. In 1985, he joined the Basic Research Laboratories, NTT (Nippon Telegraph and Telephone Corp.). Since 1992, he has been with the Department of Planning & Management of Science, Nagaoka University of Technology. His research interests include information theory, statistics, and queueing theory. Dr. Nakagawa is a member of the Institute of Electrical and Electronics Engineers and the Society of Information Theory and its Application.